TECHNICAL MEMORANDUM
January 30, 2007

GUIDELINES FOR DATA REDUCTION FOR TEM RESULTS FOR LIBBY AMPHIBOLE IN AIR AND DUST SAMPLES AT THE LIBBY SUPERFUND SITE

1.0 INTRODUCTION

When people are exposed to a chemical contaminant in an environmental medium, the long-term average level of exposure is usually assumed to be proportional to the average concentration in the area where exposure occurs. The area where exposure occurs is usually referred to as the Exposure Point, and the average concentration within the area is referred to as the Exposure Point Concentration (EPC). Typically, the EPC is estimated based on a set of measured values of the medium collected from the Exposure Point. However, the simple average of the measured values is only an estimate of the true mean, and the true mean may be either higher or lower. This is because the measured value in each sample may not be identical to the true concentration of each sample ("measurement error"), and because the set of samples collected from the Exposure Point are only a random subset of the entire set of values that occur in the Exposure Area ("sampling variability").

The USEPA has derived standard methods for computing the EPC at Superfund sites (USEPA 1989, 1992, 2002, 2004). In brief, the EPC is usually defined as the 95% upper confidence limit (UCL) on the mean, computed using appropriate statistical techniques. This approach helps minimize the likelihood that exposure and risk calculations performed for an exposure area will underestimate the true risk in that exposure unit. In some cases, especially when data are limited, the 95% UCL may substantially exceed the highest value observed, and in this situation, EPA recommends that the maximum detected value, rather than the 95% UCL, be used as the EPC.

However, existing guidance for computation of EPCs was developed mainly for use with chemical contaminants that are measured using traditional "wet chemistry" methods. At the Libby Superfund Site, the contaminant of chief concern is a form of asbestos referred to as Libby Amphibole (LA), and asbestos is measured using microscopic rather than chemical techniques. Because of this, there are several aspects of the procedure for computing exposure point concentrations of LA values that differ from the approaches used for other chemicals.

This document reviews these asbestos-specific issues associated with the derivation of concentration estimates for LA for use in exposure and risk calculations, and identifies the recommended strategy for data reduction at the Libby Superfund Site.

Note that the methods and conclusions presented in this document should not necessarily be assumed to apply to other forms of asbestos or to data from other sites.

2.0 BASIC EQUATIONS

When samples of air or dust are analyzed for asbestos using microscopic techniques such as transmission electron microscopy (TEM), the results are expressed in terms of the number of asbestos structures observed (N) divided by the total amount of sample examined (e.g., cc of air for air samples, cm^2 of surface for dust samples).

C(air) (s/cc) = N / Volume of air (cc)
C(dust)[1] (s/cm^2) = N / Area of surface (cm^2)

For convenience, analytical sensitivity (S) is defined the inverse of volume or area examined:

S(air) = 1 / Volume of air (cc)
S(dust) = 1 / Area of surface (cm^2)

Thus, concentration (both air and dust) is usually calculated as:

$$C = N \cdot S$$

Note that sensitivity is a function only of the amount of sample examined, not of the amount of asbestos in the sample:

$$S(air)\,(cc)^{-1} = \frac{EFA}{GO \cdot Ago \cdot V}$$

$$S(air)\,(cm^2)^{-1} = \frac{EFA}{GO \cdot Ago \cdot Area}$$

where:

[1] Measures of the amount of LA in dust (s/cm^2) are more accurately thought of as loading rather than concentration, but for convenience, dust values are referred to as concentration in this document.

EFA	=	Effective filter area (mm^2)
GO	=	Number of TEM grid openings examined
Ago	=	Area of each TEM grid opening (mm^2)
V	=	Volume of air passed through the filter (cc)
Area	=	Area of surface vacuumed onto the filter (cm^2)

In principle, the sensitivity for any sample of air or dust can be reduced to any value desired simply by examining more of the sample (i.e., by counting more grid openings), and there is no inherent limit imposed by the instrument.

3.0 COMBINING RESULTS FROM MULTIPLE ANALYSES

3.1 Pooling Results for Multiple Analyses of a Single Sample (Same Analytical Method)

In the event that a single air sample has been analyzed more than one time (e.g., by initially counting 10 grid openings and subsequently counting 40 additional grid openings in order to improve the sensitivity), assuming that the same analytical method (i.e., preparation, counting rules, etc.) was used in both analyses, the results may be combined by "pooling" the total counts observed and the total volume examined, as follows:

C(air) (s/cc) = (Total structures observed) / (Total volume examined)

$$= \frac{\sum N}{\sum V} = \frac{\sum N}{\sum (1/S)}$$

The equation for pooling dust concentration (loading) values is entirely analogous, except that results are expressed in units of s/cm^2 rather than s/cc.

3.2 Combining Results Across Multiple Samples (Same Analytical Method)

In cases where multiple samples (n) have been collected from a particular medium (e.g., air or dust) at some specified exposure location, if it is assumed that the concentration of LA in that medium at that location is approximately homogeneous, and if all of the samples were evaluated using the same method (i.e., same preparation steps, counting protocols, etc.), the results may be pooled across samples as described above for multiple analyses of the same sample. For example, this approach may be appropriate for combining results for multiple samples of indoor air collected

within the same room or building, since air within a room or building is often assumed to be approximately homogeneous due to mixing and circulation.

In cases where it is not appropriate to assume that the concentration in a medium is homogeneous, but may vary from sample to sample as a function of time or space, then the best estimate of the mean concentration is obtained by computing the concentration for each sample, and then averaging across all of the n samples from that location:

$$\overline{C}(air) = \frac{\sum C(i)}{n}$$

For example, in a location where indoor dust levels were clearly different between different floors of a building, it might be appropriate to pool all dust values from each floor, then average across the different floors in the building.

When results from multiple samples are combined by averaging or pooling, the scientific rationale for the approach selected must be provided as part of the evaluation.

4.0 COMBINING RESULTS BETWEEN ISO 10312 AND AHERA COUNTING RULES

Over the course of the investigation at the Libby Site, two separate sets of counting rules have been employed for TEM analysis of samples of air and dust: ISO 10312 (ISO 1995) and AHERA (AHERA 1986)[2]. Thus, an issue of general importance is the degree to which results from different counting methods are comparable, and whether it is appropriate to combine results obtained using different counting rules. This issue applies both to individual samples that have been analyzed by more than one approach, as well as to the comparison and combination of results across different samples that have been evaluated by different methods.

The chief difference between ISO and AHERA counting rules is that some types of complex structures (e.g., disperse clusters and matrices) are counted as single particles in AHERA, while they are usually separated out into component substructures in ISO. Because of these differences between the counting rules, analyses of samples by the AHERA method may tend to yield lower concentration values than by ISO. However, there are several lines of evidence which suggest that, at the Libby Site, differences between the ISO and AHERA counting methods are likely to be minor.

[2] In most cases, only particles with an aspect ratio of 5:1 or greater were recorded, as specified in the methods. For some projects (e.g., Phase 2), counting rules were revised to require recording and counting of particles with an aspect ratio of 3:1 or greater, as discussed in Laboratory Modifications LB-00016 and LB-000031. This variation in counting rules over time is not believed to be a source of substantial uncertainty in the comparison of results across methods.

Direct Comparison of Paired AHERA and ISO Analyses

As of November 6, 2006, a total of 1,869 samples of air have been analyzed by both AHERA and by ISO counting rules. Appendix A (an electronic Excel spreadsheet) provides the data for all of these samples. The results are summarized below:

Prep Method	Total Pairs	Results of Paired Comparison				
		Both Non-Detect	One or Both Detects	AHERA<ISO	AHERA=ISO	AHERA > ISO
Direct	1837	1334	503	23	467	13
Indirect	32	22	10	0	10	0
Either	1869	1356	513	23	477	13

As seen, most (1837 out of 1869) of the samples were direct preparations. Of these, most (1334) were non-detect (ND) by both ISO and AHERA. Of those that were detects by one or both methods, most (467 out of 503) were not statistically different. Of those that were different, there were slightly more pairs where AHERA was lower than ISO (23 out of 503) than where AHERA was higher than ISO (13 out of 503). For 32 samples with an indirect preparation, there were 10 samples with a detection by one or both methods, and 10 out of 10 of these were not statistically different. These results support the conclusion that any differences between AHERA and ISO counting rules for LA structures in air are likely to be minimal for samples from the Libby Site.

For dust, only one paired ISO/AHERA result was located. The results for this sample were not statistically different from each other. However, it is clear that one sample is not sufficient to support a meaningful comparison.

Frequency of Particles Counted Differently

As noted above, the main difference between ISO and AHERA counting rules is that, in ISO, some complex structures are broken down into their component elements, while in AHERA most complex particles are counted as a single structure. However, at the Libby Site, the frequency with which LA occurs in complex structures that would be counted differently by AHERA than by ISO (disperse clusters and matrices with more than 1 substructure) is relatively low. Results of a query performed on November 6, 2006, are provided below:

Medium	Preparation Method	Total Number of LA Particles	% of ISO Particles that would be Counted	Average Number of Countable Sub-	Magnitude of the

		Observed	Differently by AHERA	structures per Primary Particle	Difference
Air	Direct	5310	9.3%	2.7	16%
	Indirect	962	6.4%	2.7	11%
Dust	Indirect	2436	5.2%	2.6	8%

As seen, only about 5%-9% of all LA structures identified in ISO analyses were of a particle type that would have been counted differently by AHERA than ISO. For this subset of complex particles, the average number of countable fibers or structures delineated by ISO was about 2.7 per complex structure. Based on this, the expected average magnitude of the difference between ISO and AHERA counts for air samples (direct preparation) is calculated as:

$$\% \text{ Difference} = 9.3\% \cdot (2.7 - 1.0) = 16\%$$

Values for indirect air samples and dust samples are similar (8%-11%)

Based on this evaluation approach, differences between AHERA and ISO results are expected to be generally small (< 20%) for both air and dust samples from the Libby Site.

Default Guideline

Based on the weight of evidence from the two lines of evidence cited above, it is concluded that differences in LA particle counts between ISO and AHERA will generally be small, both for air and dust, and that the benefit of combining the results across counting methods (decreased statistical uncertainty due to larger sample size) will generally outweigh any minor bias or error that might be introduced by combining the results. Thus, the default guideline is that results for ISO and AHERA analyses may be compared and combined without adjustment, both within and between samples. However, each data user must consider the pros and cons of this approach for their intended data use, and document such considerations and supporting rationale when evaluating the data.

5.0 COMBINING AIR RESULTS FROM DIRECT AND INDIRECT PREPARATION

Most air samples at the Libby Site are analyzed using a direct preparation, but if the sample is overloaded with particulate matter, an indirect preparation may be required. The primary issue associated with combining or comparing the results of direct and indirect analyses is that the steps used in the indirect preparation (e.g., suspension of the sample in water, usually accompanied by sonication) may cause some asbestos structures to disaggregate into smaller particles, thereby increasing the number of countable structures.

A number of studies have been performed at other locations to investigate the effect of indirect preparation on estimates of asbestos concentration. Useful reviews are provided in HEI-AR (1991) and Breysse (1991). In general, the available data indicate that the magnitude of any difference between a direct and an indirect preparation for a sample depends on the nature of the asbestos (chrysotile vs. amphibole), and the details of the indirect preparation technique, especially the duration and energy of sonication. For chrysotile, the magnitude of the increase in estimated concentration (s/cc) due to indirect preparation is usually in the range of 2-100 fold (e.g., Hwang and Wang 1983, Sahle and Laszlo 1986), but may sometimes be as large as 1000-2000-fold (e.g., Kauffer at al. 1996, Chesson and Hatfield 1990). The magnitude of the difference is often larger for short particles (e.g., length < 5 um) than for longer particles (Hwang and Wang 1983, Chatfield 1985, Kauffer et al 1996). For amphiboles, differences between direct and indirect preparation are generally much smaller (less than 10-fold) than for chrysotile (e.g., Bishop et al. 1978, Sahle and Laszlo 1996).

Direct Comparison of Paired Air Samples from the Libby Site

At the Libby Site, very few air samples are analyzed by both direct and indirect preparations. As of September 5, 2005, only 16 examples of this type existed in the Libby database. Appendix B provides the data for these 16 samples. Of these 16 paired results, six were non-detect (ND) by both direct and indirect preparations. Although these ND-ND pairs rank as "agreement" between the direct and indirect preparation methods, it is more revealing to compare results for pairs in which one or more analysis identified one or more LA structures. Of these (a total of 10 pairs), nine of the 10 were not statistically different between direct and indirect preparations, and one of the 10 was statistically higher for the indirect preparation than the direct preparation.

Because there were so few paired results in the existing database, especially for results that are not ND-ND, a set of 31 air samples that had previously been analyzed by TEM (AHERA) using a direct preparation method were selected for reanalysis by TEM (AHERA) using an indirect preparation method to evaluate the potential effect of indirect preparation on the number and types of LA structures observed in air samples from the Libby Site. Appendix C provides the basic study design and describes how these samples were selected. Appendix D provides the detailed results for these 31 samples, both for the original direct analysis and the indirect reanalysis.

Table 1 summarizes the total LA results for these 31 samples, and the values are shown graphically in Figure 1 (Panel A). When compared pair-wise, indirect preparation samples were statistically higher ($p < 0.05$) than the matched direct preparation samples in 14 of 31 (45%) cases (red symbols), were statistically lower in 7 of 31 (23%) cases (green symbols), and were not statistically different in 10 of 31 (32%) of the cases (black symbols). The Wilcoxon signed rank test indicates that the paired data sets (indirect *vs.* direct) are not significantly different from each other, although

the difference is close to being significant ($p = 0.07$). As shown in Figure 1 (Panel B), if the comparison is restricted to LA structures longer than 5 um, the frequency and magnitude of the differences are diminished, but there are still a number of samples in which the indirect preparation is several times higher than the direct preparation.

Particle Type and Size Evaluation

Another way to investigate the effect of indirect preparation is to examine the frequency of complex particles observed in the 31 samples evaluated by both direct and indirect preparations:

Particle Type	Percent of Total	
	Direct	Indirect
Fiber	72%	67%
Bundle	8%	1%
Cluster	0.1%	0.0%
Matrix	21%	32%

As seen, these data are consistent with the hypothesis that indirect preparation tends to decrease the occurrence of bundles, but also indicate that there is an increase in the number of matrix particles, perhaps due to breakup of large matrix particles into smaller matrix particles during sonication.

Figure 2 compares the length and width distributions for LA particles observed in the 31 samples analyzed by both direct and indirect preparation. For length (upper panel), the distribution for LA structures observed in indirect preparations tends to be left-shifted from that for direct preparations, suggesting that indirect preparation may tend to cause breakage of some long LA fibers into shorter fibers. This difference is statistically significant (Wilcoxon Rank Sum test, $p < 0.001$). For width (lower panel), the distributions are generally similar to each other except at the high end, where structures observed in indirect preparations tends to be thicker than for direct preparation. Although the difference is significant (Wilcoxon Rank Sum test, $p = 0.050$), this observation is not expected and may be due to random variation, since it is not apparent how an indirect preparation can cause LA particles to become thicker.

Discussion

Based on the various lines of evidence described above, it is concluded that indirect preparations may alter the estimated concentration of LA particles in samples of air. There is a general tendency for LA concentrations to be increased by indirect preparation, although in some samples the concentrations are apparently decreased. In many cases, the differences are within a factor of 2-3, but some differences may be larger. The differences appear to be less important for long structures

(> 5 um) than for total LA (where structures < 5 um in length are also presented). The basis for the differences appears to be a complex interaction of multiple factors, including disaggregation of some complex structures into fibers, breaking of some long structures into shorter structures, and dispersion of large matrix particles into many smaller matrix particles. Because of these differences, inclusion of data from indirect preparations in the computation of exposure point concentrations for air may yield results that are different (usually higher, but sometimes lower) than would be obtained if the data were from direct preparation samples only.

Default Guideline

Because most (>97%) of all air samples collected in Libby have been analyzed using the direct preparation method, and because the difference between direct and indirect samples is usually within a factor of 2-3 (especially for long fibers), any error introduced into exposure and risk evaluations by use of occasional indirect data is likely to be minimal. Based on this, the default guidance for the Libby Superfund Site is that results for direct and indirect preparations or air samples may be compared and combined without adjustment, both within and between samples. However, each data user must consider the pros and cons of this approach for their intended data use, especially when the data for a particular location are based primarily on indirect data. In this case, the uncertainty associated with reliance on indirect data and the potential direction and magnitude of bias shall be discussed as part of any evaluation.

6.0 DEALING WITH SAMPLES WITH ZERO OBSERVED STRUCTURES

If zero structures are observed (N = 0) when a sample of air or dust is analyzed, this is generally referred to as a "non-detect" result. For analytes other than asbestos, EPA suggests that, when computing the mean of a set of samples, "non-detects" (i.e., samples whose concentration is below the detection limit of the analytical instrument) be evaluated by assigning a surrogate value of ½ the detection limit (USEPA 1989). By analogy, it is sometimes supposed that "non-detects" for asbestos should be evaluated by assigning a value equal to ½ the sensitivity. However, this is <u>not correct</u>. The analytical sensitivity in microscopic analyses is not analogous to a detection limit in a wet chemistry analysis, and use of ½ the sensitivity as a surrogate for asbestos non-detects may lead to a substantial overestimate of the true mean of a group of samples.

This is demonstrated in Figure 3. As seen, in cases where the analytical sensitivity is larger than the true concentration, if an asbestos non-detect is assigned a value equal to ½ the analytical sensitivity, the estimate of the mean will be biased high, with the magnitude of the error tending to increase as the ratio of sensitivity to true concentration increases. In cases where the analytical sensitivity is less than about ½ the true concentration, the magnitude of the error introduced by assigning ½ the sensitivity to non-detects becomes negligible.

There are two reasons why a non-detect in microscopy is not analogous to a non-detect in a traditional wet chemistry analysis, and should not be evaluated by assigning a value of ½ the sensitivity:

- A non-detect by a non-microscopic technique indicates that the amount of analyte present in the fraction of the sample placed into the analytical instrument was less than the detection limit, while a non-detect by a microscopic technique indicates that the amount of asbestos present in the fraction of the sample evaluated under the microscope was truly zero. Note that this statement is not inconsistent with the recognition that the observation of zero structures in some particular set of grid openings examined does not prove that there are zero structures in other grid opening that were not examined. This topic (uncertainty in the observed number of structures observed) is discussed in Section 7 (below).

- The results of a wet chemistry method yield continuous values, while results of a microscopic result yield discrete (discontinuous) values. That is, the concentration value reported in a microscopic analysis can only occur in multiples of the analytical sensitivity S (e.g., 0S, 1S, 2S, etc.). This means that when the true concentration of a sample is lower than the sensitivity, any and all detects will yield concentrations that are higher than the true value, rather than a reliable estimate of the true value. For example, consider the case where the true concentration is 0.001 s/cc, and the sensitivity is 0.010 s/cc. If this sample were analyzed 10 times, the expected result would be that about 9 of the 10 analyses would yield a count of zero, and one of the samples would yield a count of 1, which would correspond to a concentration estimate of 0.010 s/cc (10-times the truth). Only when the occasional high values are averaged with the "non-detects" does the estimate of the mean approach the true value.

This topic (the correct statistical approach for evaluating non-detect values from discontinuous count-based measurement methods) has been reviewed by EPA previously as part of the rulemaking process for microbial contamination in drinking water (USEPA 1999). (Note: measurement of pathogens in water is closely analogous to measurement of asbestos structures in air, in that the analysis is based on visual observation and yields discrete rather than continuous results). During a public workshop held on this topic in 1998, a number of statistical experts provided information on the correct methods for computing the concentration of *Cryptosporidium* in source waters of drinking water supplies, given that some (most) of the individual samples were "non-detect" (i.e., no spores of *Cryptosporidium* were observed in the sample analyzed). The expert panel emphasized that "non-detects" for *Crytosporidium* that occur in a set of water samples from a water system must be evaluated with a value of zero when computing summary statistics on the mean level of organisms present in the water.

7.0 DEALING WITH UNCERTAINTY

As noted above, all estimates of environmental concentration values are uncertain because the measured value in each sample may not be identical to the true concentration of each sample ("measurement error"), and because a random set of samples collected from an Exposure Area may not be representative of the true average in the Exposure Area ("sampling variability"). The following sections describe statistical approaches for characterizing the uncertainty in concentration values for individual samples and in the mean of multiple samples.

7.1 Uncertainty in Individual Sample Values Due to Poisson Variation

All analytical results are associated with some degree of measurement error. That is, repeat analysis of multiple independent aliquots from the same sample usually do not yield identical results. This applies equally to analysis of asbestos and traditional wet chemistry materials.

For asbestos, for a single analysis of a sample, the concentration is estimated as:

$$C = N \cdot S = N / V$$

where:

C = Concentration (f/cc)
N = Number of countable asbestos fibers observed
S = Analytical sensitivity (cc^{-1})
V = Volume of air (cc) that passed through the area of filter examined

However, because N (the number of structures observed in the area of filter examined) is a Poisson random variable, the value of C is uncertain. The probability density function (PDF) that characterizes the uncertainty around the observed concentration is given by (Box and Tiao 1992):

$$PDF(C) \sim CHISQ(2 \cdot N+1) / (2 \cdot V)$$

where:

CHISQ(υ) = Chi-squared distribution function with υ degrees of freedom

Because $N = C \cdot V$, the uncertainty distribution around the observed count of N may be expressed as:

$$PDF(N) \sim 0.5 \cdot CHISQ(2 \cdot N+1)$$

Figure 4 shows the uncertainty PDFs for three hypothetical samples with counts of 0, 3, or 10 structures. As seen, for a sample with a count of zero (Panel A), the uncertainty distribution includes zero (it is the most likely value), but the distribution is right skewed and extends out to include plausible values as high as 2 or even higher. As the observed count becomes larger, the uncertainty distribution becomes more nearly centered on the sample observation, and tends to become more symmetric (Panels B and C).

Note that if a single sample has been analyzed more that one time and the results are pooled (see Section 3.1), this same approach may be used to characterize the uncertainty in the pooled analysis.

At the Libby Site, each analytical measurements should be reported with a description of the statistical uncertainty around the measurement. The statistic recommended for normal reporting is the two-sided 90% CI (5^{th} percentile to the 95^{th} percentile). This interval will include the true concentration in approximately 90% of all samples, and there is 95% confidence that the true concentration is less than or equal to the upper bound. However, other confidence intervals may be presented when this is considered to be important in proper characterization and interpretation of the data.

7.2 Uncertainty in the Mean of Multiple Samples with Poisson Variation

In cases where multiple samples have been collected from an exposure area and pooling is not considered to be appropriate, uncertainty in the mean concentration that is attributable to random Poisson variation in the analytical count can be characterized using Monte Carlo simulation, as follows:

Step 1. For each sample, specify the uncertainty distribution around the observed concentration as described above:

$$PDF(C_i) \sim CHISQ(2 \cdot N_i+1) / (2 \cdot V_i)$$

Step 2. Using an appropriate computer software application, draw one random value from each sample and compute the mean.

Step 3. Repeat Step 2 many times. Select the 95^{th} percentile of the means as a conservative estimate of the true sample mean.

It is important to stress that this approach evaluates only the uncertainty arising from random Poisson variation in count, and does not include uncertainty due to variation in concentration over time and space. Thus, the UCL-based on Monte Carlo simulation will usually underestimate the combined UCL, except in special cases where there is very little spatial or temporal variability between samples. An approach for characterizing uncertainty from both sources is presented in the next section.

7.3 Uncertainty in the Mean of Multiple Samples Due to Poisson Measurement Error Combined with Spatial or Temporal Variability

When a set of samples is collected from an exposure area in which concentration varies over space or time, the resulting data values include the between-sample variability that arises from both analytical measurement error in individual samples and from between-sample temporal or spatial variability. The mathematical procedure for computing the 95% UCL of the mean for a data set depends on the attributes of the data set. In some cases, the distribution of values may be well approximated by a parametric distribution (normal, lognormal, gamma), and the UCL may be computed using an equation appropriate for the underlying distribution. In other cases, the data may not be well approximated by a parametric distribution, and estimation of the UCL may achieved through non-parametric procedures (e.g., Chebychev inequality, bootstrapping).

7.3.1 Attributes of Data Sets from Libby

Summary statistics for five example data sets from Libby are presented below:

Set	Description	N	Detect. Freq.	Mean Sens. (cc^{-1})	Mean Conc. (s/cc)	Min Detect (s/cc)	Max Detect (s/cc)
1	Ambient air (all)[3]	404	15%	0.00248	0.00065	0.00020	0.03327
2	Ambient air (re-analyzed)	33	45%	0.00010	0.00021	0.00007	0.00241
3	Personal air (disturbed soil)	67	54%	0.00270	0.0507	0.00096	1.343
4	Personal indoor air	21	86%	0.00036	0.00119	0.00007	0.0066
5	Stationary indoor air	31	65%	0.00007	0.00017	0.00006	0.00080

As seen, based on these 5 examples, some data sets from Libby may be characterized by relatively high non-detect frequencies, and most appear to be characterized by a wide range of values (usually 2-3 orders of magnitude or more between the minimum and maximum detect). In general, statistical evaluation is especially difficult for data sets that are moderately to heavily censored and are highly variable.

[3] Represents all ambient air samples collected prior October 2006, when the Outdoor Ambient Air Monitoring Program was implemented.

Figure 5 presents log-probability plots for the 5 data sets evaluated above. Detects are shown as solid symbols, while non-detects are shown as open symbols[4]. As seen, the detects tend to be moderately well-fit by a straight line, indicating that an assumption of lognormal variation between samples may be reasonable in most cases. Assuming lognormality, and ignoring the contribution of Poisson measurement error, the parameters of the lognormal distribution may be crudely estimated from the parameters of the best fit line through the detects (Gilbert 1987):

σ = slope
μ = intercept

The corresponding linear space parameters are computed from the log-space parameters as follows:

$$\text{mean} = \exp(\mu + 0.5 \cdot \sigma^2)$$
$$\text{GSD} = \exp(\sigma)$$

As seen in Figure 5, the value of GSD (an indication of the degree of between-sample variability) is quite large in most cases, with estimated values above 5 in four out of the five data sets. The GSD of 35.8 for data set 3 (activity based samples) is especially high because the data set is composed of samples that span a wide range of soil levels (ranging from clean fill to soil with LA above 1%).

7.3.2 Evaluation of ProUCL for Calculating the UCL

EPA has developed a software package (ProUCL) and guidance for how to compute a UCL value for a given data set (USEPA 2002, 2004). In brief, the data set is entered into ProUCL and the software computes the UCL based on a series of alternative parametric and non-parametric methods. Then, based on an analysis of the attributes of the data set (magnitude of between-sample variability, number of samples, degree of skewness), the software recommends one of the UCL values as being most appropriate.

In order to evaluate the performance of ProUCL for data sets similar to the examples described above, a series of synthetic data sets of 15 samples each were generated using Monte Carlo simulation, with varying levels of measurement error (modeled as a Poisson random variable) and/or spatial variability (modeled as a lognormal random variable). For each random data set, the 95% UCL was computed using the Chebychev inequality method, since this was the equation selected by ProUCL for each of the five example data sets:

[4] The plotting position for non-detects is arbitrary

$$UCL \leq mean + \sqrt{(1/\alpha) - 1} \cdot stdev / \sqrt{n}$$

The resulting UCL value was scored either as "pass" (UCL ≥ true mean) or as "fail" (UCL < true mean). The results are summarized in Figure 6. As seen, the failure rate of the Chebychev UCL increases as the degree of temporal/spatial variability increases, being ≤ 5% only when the spatial/temporal variability is relatively low (GSD = 2.3). The magnitude of the measurement error and the detection frequency appears to have relatively little effect.

Because it appears likely that most data sets from Libby will tend to have GSD values larger than 2-3, it is concluded that ProUCL is unlikely to yield UCL values that can be relied on to have a high probability of exceeding the true (but unknown) mean. For this reason, ProUCL will not be used as the main statistical tool for estimating UCLs for data sets at Libby.

7.3.3 Other Methods for Estimating the UCL

a. Fitting to a Poisson-Lognormal (PLN) Distribution

Based on the log-probability plots shown in Figure 5, it appears that an assumption of lognormality may be reasonable for air data sets from Libby (at least for the 5 examples provided). If so, then the approach for computing the 95% UCL may be based on the equations appropriate for a lognormal distribution, as described in USEPA (1992):

$$UCL = \exp[\mu + 0.5 \cdot \sigma^2 + \sigma \cdot H / \text{sqrt}(n-1)]:$$

The values of μ and σ may be estimated by fitting the data set to a Poisson-lognormal (PLN) distribution, as described by Haas et al (1999). The mathematical details are provided in Attachment 3.

Figure 7 shows the fraction of all 95% UCL values generated by this approach that are greater than or equal to the true mean as a function of the underlying variability (GSD = 5 or 10) and as a function of the ratio of the true concentration to the analytical sensitivity. As seen, the UCL coverage (the fraction of all UCLs that equal or exceed the true mean) tends to increase as the average number of counts in the sample increases, approaching a maximum of slightly less than 95% when the average count is 10 or above. Because the goal is to achieve 95% coverage, the method was modified to compute the 97.5% UCL rather than the 95% UCL, as shown in Figure 8. As seen, this yields a coverage of approximately 95% when the average number of counts is 10 or above. However, coverage tends to drop off for cases where the average count is less than 10.

b. Two-Step Bootstrap Method

A second method for estimating the 95% UCL of a data set that combines Poisson measurement error with an underlying lognormal (or other) sampling distribution is bootstrapping. In bootstrapping, the observed data set is used to generate many alternative data sets by random sampling with replacement, and the 95th percentile of the bootstrap means is taken as an estimate of the 95% UCL of the sample mean. However, this simple (1-step) method does not account for the variation due to Poisson measurement error. To account for this added source of variation, a second step is added to the method. In this approach, the first step is generation of a non-parametric bootstrap sample, as above. Then, for each observation in the bootstrap sample, a new concentration value is generated by making a random draw from the uncertainty distribution around the observation (concentration value), given by CHISQ(2N+1) / (2V).

Figure 9 shows the UCL coverage for this method. As seen, UCL coverage tends to be lower than 95% when the average count begins to exceed 1.0, but is high when the average count is low. Note that this pattern is essentially the opposite of what is seen for the PLN-Land approach (Figures 7 and 8). That is, data sets that tend to have low UCL coverage using the PLN-Land method will tend to have high UCL coverage using the 2-step bootstrap method.

c. Side-by-Side Comparison of Method Performance

In order to compare the effectiveness of these two approaches for computing UCL values, random data sets of varying size and with varying degrees of sampling variability (controlled by setting the size of the GSD) and Poisson measurement error (controlled by setting the ratio of the true mean concentration and the analytical sensitivity, C/S) were generated using Monte Carlo simulation as described above, and the UCL for each data set was generated using each method. The results were assessed by dividing each UCL by the true mean of the lognormal distribution from which the data set was drawn, and examining the cumulative distribution function (CDF) of the resulting ratio. The "ideal" attributes of such a CDF is there is high coverage (i.e., the percent of all UCL values that are equal to 1.0 is $\geq$ 95%), and there is a relatively low fraction (e.g., $\leq$ 30%) of all values larger than about 3.0. If the CDF of values were to achieve this pattern, then the probability of a Type I error would be $\leq$ 5%, and the probability of a Type II error would likely be within acceptable bounds.

The results are shown in Figure 10. This figures compares the CDFs of the UCL/mean ratio for the PLN-Land method to the two-step bootstrap method for 16 different combinations of samples size (10, 25, 50 or 100) and four different detection frequencies (0-20%, 21-40%, 41-60%, and > 60%). Inspection of these panels reveals the following main points:

- For low detection frequencies (< 20%), the two-step bootstrap method yields UCLs with higher coverage than the PLN-Land method, independent of sample size. In addition, the bootstrap CDFs are narrower and steeper than for the PLN-Land method, so there is a lower frequency of excessively high UCL values. Based on this, the two-step bootstrap method is identified as the preferred method when detection frequencies are < 20%.
- For detection frequency of 21-40%, the UCL coverage is generally similar between the two-step bootstrap method and the PLN-Land method. However, as noted above, the bootstrap CDFs tend to be steeper than for the PLN-Land approach, limiting the number of large values. Based on this, the two-step bootstrap method is identified as the preferred method when detection frequencies are 21-40%.
- For detection frequencies greater than 40%, the UCL coverage is better for the PLN-Land method than the two-step bootstrap. Thus, even though the bootstrap curves tend to be steeper, the PLN-land method is preferred because of the better UCL coverage.
- For any specified detection frequency (which is related to the ratio of concentration and analytical sensitivity), increasing the sample size tends to increase the UCL coverage and decrease the fraction of high values, especially for the PLN-Land method. At high sample size and high detection frequency, the two methods tend to approach each other.

7.5 Default Guideline

Because air data sets from Libby often have a relatively high non-detect frequency and usually have a relatively high degree of skewness, calculation of 95% UCL values is difficult. Tests using EPA's standard software package for computing UCLs (ProUCL) indicate that this software package does not handle this type of data set well, this method will not be relied on as the primary analysis tool at the Libby site.

The best alternative approach available at present appears to depend on the attributes of the data set. For data sets with a detection frequency of $\leq$ 40%, the 2-step bootstrap method appears to yield UCLs with good coverage and a relatively low probability of excessively high values. For data sets with higher detection frequencies, the UCL coverage of the bootstrap method begins to decrease and the PLN-Land method becomes preferred.

To facilitate implementation of this approach, an executable program has been developed that accepts as input a site-specific data set (n paired values of count and volume), and computes the UCL by the 2-step bootstrap method and also computes the values of mu and alpha derived by PLN fitting method. These parameters may be used to compute the PLN-Land UCL externally. This tool is provided as Attachment 4.

Region 8 will continue to investigate methods for estimating the UCL of the mean for sample sets similar to those expected to occur in Libby, and may revised the recommended approach when improved statistical techniques are identified.

8.0 CONCLUSIONS

At the Libby Superfund Site, available data support the following conclusions regarding the estimation of LA in samples of air or dust:

- Estimates of LA concentration in air do not differ substantially when measured using ISO 10312 and AHERA counting rules, and results may be compared and combined for data evaluation. However, each data user must consider the pros and cons of this approach for their intended data use, and document such considerations and supporting rationale when evaluating the data.

- Paired data that compare ISO and AHERA results are not available for samples of dust, but the frequency of particles in dust samples that would be counted differently by the two methods is low. Thus, it is considered likely that dust samples analyzed by ISO and AHERA will also generally be similar, and hence they may also be combined across methods. However, each data user must consider the pros and cons of this approach for their intended data use, and document such considerations and supporting rationale when evaluating the data.

- In some (but not all) samples of air, estimates of LA concentrations may be several fold higher when measured using an indirect preparation compared to a direct preparation. Thus, use of the indirect results may tend to overestimate exposure and risk estimates in some cases. Because of the low frequency of indirect preparations for air samples at the Libby Site, this is likely to be a minor source of uncertainty in most cases, but should be identified as a source of uncertainty whenever exposure point concentration values are based primarily on indirect samples.

- When computing the best estimate of the arithmetic mean concentration of asbestos for an exposure point, all non-detect values must be evaluated using a concentration of zero. This is in contrast to the approach used for most other chemicals, where ½ the detection limit is assigned to non-detects. However, the analytical sensitivity in microscopic analyses for asbestos is not analogous to a detection limit in a wet chemistry analysis, and use of ½ the sensitivity as a surrogate for asbestos non-detects may lead to a substantial overestimate of the true mean of a group of samples.

- Individual sample results should be accompanied by a characterization of the uncertainty in the values. The 2-sided 90% confidence interval is recommended for most cases. However, other confidence intervals and supporting justification for its use may be presented when it is considered to be important in proper characterization and interpretation of the data.

- Uncertainty in the mean concentration of a data set arises from the combined effect of analytical measurement error and temporal or spatial sampling variability. The best method for estimating the UCL of the mean depends on the attributes of data set. Based on analyses performed to date, the-two-step bootstrap method is best for samples with low detection frequency (<40%), while the PLN-Land method is preferred for data sets with higher detection frequencies. This approach for estimating UCL values may be refined in the future.

It is important to stress that these conclusions and recommendations may not apply to other forms of asbestos or to data from other sites.

9.0 REFERENCES

AHERA. 1986. Asbestos Hazardous Emergency Response Act . Title 20, Chapter 52, Sec. 4011. Public Law 99-519.

Bishop K, Ring S, Suchanek R, Gray D. 1978. Preparation Losses and Size Alterations for Fibrous Mineral Samples. Scanning Electron Microsc. I:207.

Box GEO, Tiao GC. 1992. Bayesian Inference in Statistical Analysis. John Wiley and Sons.

Breysse PN. 1991. Electron Microscopic Analysis of Airborne Asbestos Fibers. Crit. Rev. Analyt. Chem. 22:201-227.

Chatfield EJ. 1985. Limitations of Precision and Accuracy in Analytical Techniques Based on Fiber Counting, in *A Workshop on Asbestos Fiber Measurements in Building Atmospheres.* Chatfield EJ (Ed). Ontario Research Foundation, Ontario, Canada. 115.

Chesson J, Hatfield J. 1990. Comparison of Airborne Asbestos Levels Determined by Transmission Electron Microscopy (TEM) Using Direct and Indirect Transfer Techniques. U.S. Environmental Protection Agency, Exposure Evaluation Division, Office of Toxic Substances, Office of Pesticides and Toxic Substances.

Gilbert RO. 1987. Statistical Methods for Environmental Pollution Monitoring. Van Nostrand Reinhold, New York.

Gilliom RJ. Helsel DR. 1986. Estimation of Distributional Parameters for Censored Trace Level Water Quality Data: Part 1. Estimation Techniques. Water Resources Research 22:135-146.

Haas CN, Rose JB, Gerba CP. 1999. Quantitative Microbial Risk Assessment. John Wiley and Sons, New York.

HEI-AR. 1991. Asbestos in Public and Commercial Buildings: A Literature Review and Synthesis of Current Knowledge. Health Effects Institute – Asbestos Research. Cambridge, Massachusetts.

Hwang CY, Wang ZM. 1983. Comparison of Methods of Assessing Asbestos Fibre Concentrations. Arch Environ. Hlth. 38:5-10.

ISO. 1995. International Organization for Standardization (ISO). Ambient Air – Determination of Asbestos Fibres – Direct-Transfer Transmission Electron Microscopy Method. ISO 10312:1995(E).

Kauffer E, Billon-Galland MA, Vigneron JC, Veissiere S, Brochard P. 1996. Effect of Preparation Methods on the Assessment of Airborne Concentrations of Asbestos Fibres by Transmission Electron Microscopy. Ann. Occup. Hyg. 40:321-330.

Sahle W, Laszlo I. 1996. Airborne Inorganic Fibre Monitoring by Transmission Electron Microscope (TEM): Comparison of Direct and Indirect Sample Transfer Methods. Ann. Occup. Hyg. 40:29-44.

USEPA. 1989. Risk Assessment Guidance for Superfund. Volume I. Human Health Evaluation Manual (Part A). U.S. Environmental Protection Agency, Office of Emergency and Remedial Response. EPA/540/1-89/002.

USEPA. 1992. Supplemental Guidance to RAGS: Calculating the Concentration Term. U.S. Environmental Protection Agency, Office of Solid Waste and Emergency Response. Publication 9285.7-081.

USEPA. 1999. M/DBP Stakeholder Meeting Statistics Workshop Meeting Summary: November 19, 1998, Governor's House, Washington DC. Final. Report prepared for U.S. Environmental Protection Agency, Office of Ground Water and Drinking Water by RESOLVE, Washington, DC, and SAIC, McLean, VA, EPA Contract No. 68-C6-0059.

USEPA. 2002. Calculating Upper Confidence Limits for Exposure Point Concentrations at Hazardous Waste Sites. U.S. Environmental Protection Agency, Office of Emergency and Remedial Response. OSWER 9285.6-10. December.

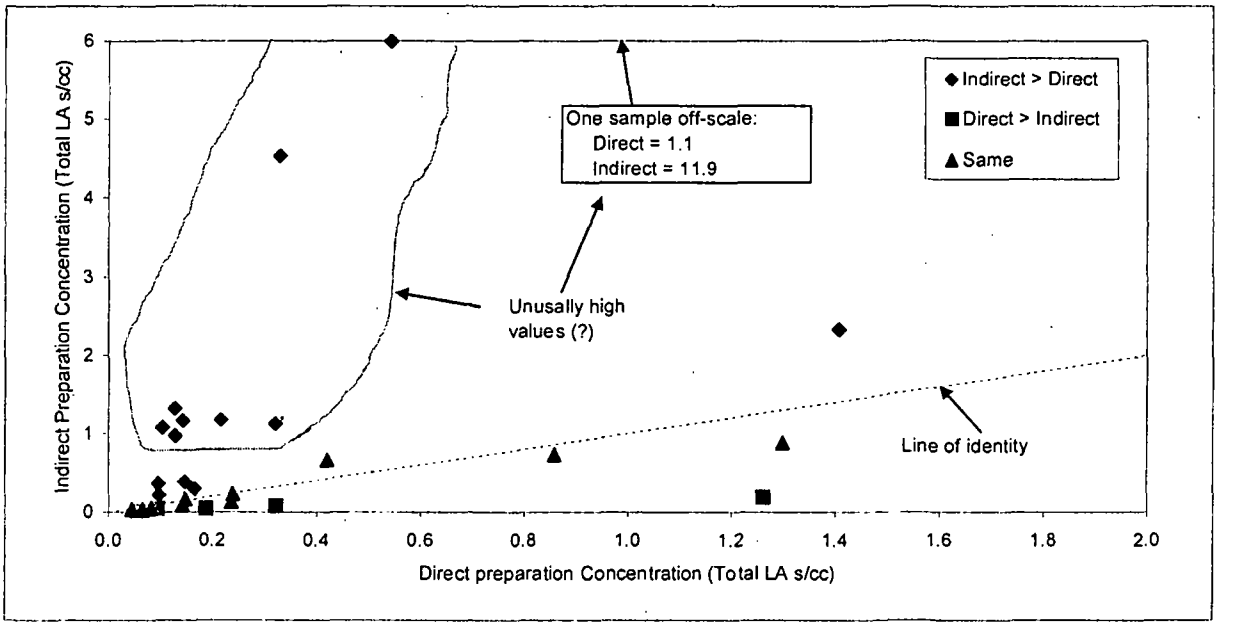
USEPA. 2004. ProUCL Version 3.0 User Guide. Report prepared for USEPA by Lockheed Martin Environmental Services. EPA/600/R04/079. April 2004.

TABLE 1
SUMMARY OF 31 AIR SAMPLES ANALYZED BY DIRECT AND INDIRECT PREPARATION

Sample	Index ID	Sample Date	Property Description	Location	Land Use	Location	Type	Personal Activity Description	LA Air Conc (s/cc)		Ratio of Conc	Statistical test
									Direct	Indirect		
1	1R-04986	5/24/01	Rainy Creek Rd	Haul Rd loop	Industrial	Outdoor	Stationary		0.096	0.363	3.78	Indirect > Direct
2	1R-08838	9/5/01	KDC Bluffs	Property	Residential	Outdoor	Personal	Labor - water hose	0.097	0.221	2.28	Indirect > Direct
3	1R-14215	7/23/02	KDC Flyway	Property	Industrial	Outdoor	Personal	Laying Curlex	0.147	0.384	2.62	Indirect > Direct
4	1R-14387	8/12/02	Rainy Creek Rd	Road	Industrial	Outdoor	Personal	Water Hose Operator	0.143	1.164	8.13	Indirect > Direct
5	1R-14413	8/20/02	Rainy Creek Rd	Road	Industrial	Outdoor	Personal	Decon	0.216	1.175	5.44	Indirect > Direct
6	1R-14508	8/12/02	Rainy Creek Rd	Road	Industrial	Outdoor	Personal	Decon	0.166	0.298	1.80	Indirect > Direct
7	1R-14528	8/16/02	Rainy Creek Rd	Road	Industrial	Outdoor	Personal	Laborer	1.102	11.893	10.79	Indirect > Direct
8	1R-14660	8/20/02	Rainy Creek Rd	Road	Industrial	Outdoor	Personal	Decon	0.128	1.321	10.36	Indirect > Direct
9	1R-14851	8/28/02	Rainy Creek Rd	Road	Industrial	Outdoor	Personal	Decon	0.104	1.079	10.36	Indirect > Direct
10	1R-14909	9/5/02	Rainy Creek Rd	Road	Industrial	Outdoor	Personal	Laborer	0.321	1.124	3.50	Indirect > Direct
11	1R-15074	9/13/02	Rainy Creek Rd	Road	Industrial	Outdoor	Personal	Upper decon	0.128	0.971	7.56	Indirect > Direct
12	1R-21599	7/15/03	Rainy Creek Rd	Road	Industrial	Outdoor	Personal	Truck Driver (Water)	0.329	4.535	13.80	Indirect > Direct
13	1R-21808	7/30/03	Rainy Creek Rd	Mine	Industrial	Outdoor	Personal	Operator-upper dozer	0.543	5.994	11.04	Indirect > Direct
14	1R-29025	5/17/05	1511 Gallatin Ave	Attic	Residential	Indoor	Personal	Bulk VCI Removal	1.408	2.332	1.66	Indirect > Direct
15	1R-04987	5/24/01	Rainy Creek Rd	Adj. #19 Terrance	Industrial	Outdoor	Stationary		0.045	0.033	0.74	Not different
16	1R-05349	6/19/01	Screening Plant	Auto	Residential	Outdoor	Stationary		0.065	0.038	0.59	Not different
17	1R-05586	6/27/01	Screening Plant	Property	Residential	Outdoor	Personal	Drive - Volvo #11	0.082	0.053	0.65	Not different
18	1R-14553	8/16/02	156 S. Central Rd	House	Residential	Indoor	Personal	Vacuuming attic	0.421	0.664	1.58	Not different
19	1R-14217	7/23/02	KDC Flyway	Property	Industrial	Outdoor	Personal	Laying Curlex	0.236	0.130	0.55	Not different
20	1R-23338	10/1/03	Rainy Creek Rd	Road	Industrial	Outdoor	Personal	Upper Dozer	1.299	0.882	0.68	Not different
21	1R-07741	8/16/01	Screening Plant Flyway	Property	Mine	Outdoor	Personal	Operate - Bulldozer, upper level	0.147	0.168	1.15	Not different
22	1R-14725	8/22/02	Rainy Creek Rd	Road	Industrial	Outdoor	Personal	Decon	0.141	0.088	0.62	Not different
23	1R-14733	8/27/02	Rainy Creek Rd	Road	Industrial	Outdoor	Personal	Deconning Truck	0.239	0.235	0.98	Not different
24	1R-29026	5/17/05	1511 Gallatin Ave	Attic	Residential	Indoor	Personal	Bulk VCI Removal	0.859	0.732	0.85	Not different
25	1R-14416	8/19/02	Rainy Creek Rd	Road	Industrial	Outdoor	Personal	Decon	0.187	0.044	0.24	Direct > Indirect
26	1R-14570	8/19/02	156 S. Central Rd	House	Residential	Indoor	Personal	Cleaning attic	0.094	0.044	0.47	Direct > Indirect
27	1R-21585	7/14/03	Rainy Creek Bank	Road	Residential	Indoor	Personal	Water Truck Driver	0.321	0.081	0.25	Direct > Indirect
28	1R-26146	8/13/04	Rainy Creek Rd - S Frontage	Property	Residential	Outdoor	Personal	Water Hose Operator	1.262	0.188	0.15	Direct > Indirect
29	1R-28513	1/7/05	4000 Pipe Creek Rd	Property	Commercial	Outdoor	Stationary		0.055	0.000	0.00	Direct > Indirect
30	1R-31263	6/22/05	105 W. 2nd St	Parking Lot	Residential	Outdoor	Stationary		0.056	0.004	0.08	Direct > Indirect
31	1R-31264	6/22/05	105 W. 2nd St	Parking Lot	Residential	Outdoor	Stationary		0.064	0.008	0.13	Direct > Indirect

FIGURE 1
COMPARISON OF DIRECT AND INDIRECT TEM RESULTS FOR 31 AIR SAMPLES FROM LIBBY

Panel A: Total LA



Panel B: Length > 5 um

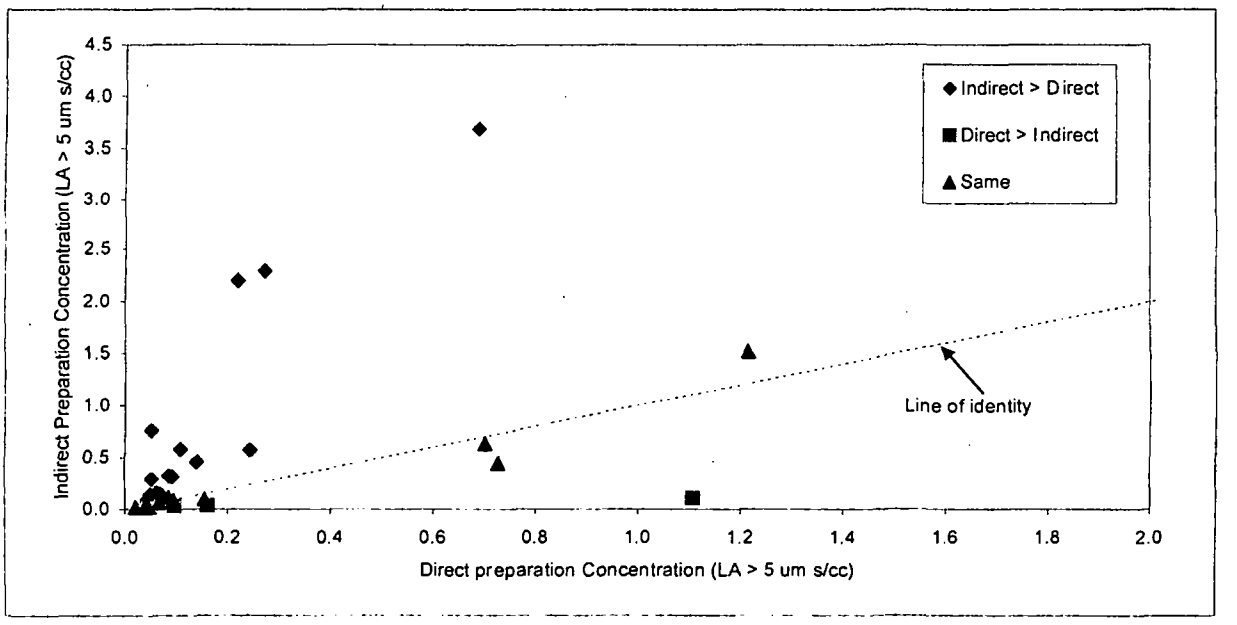
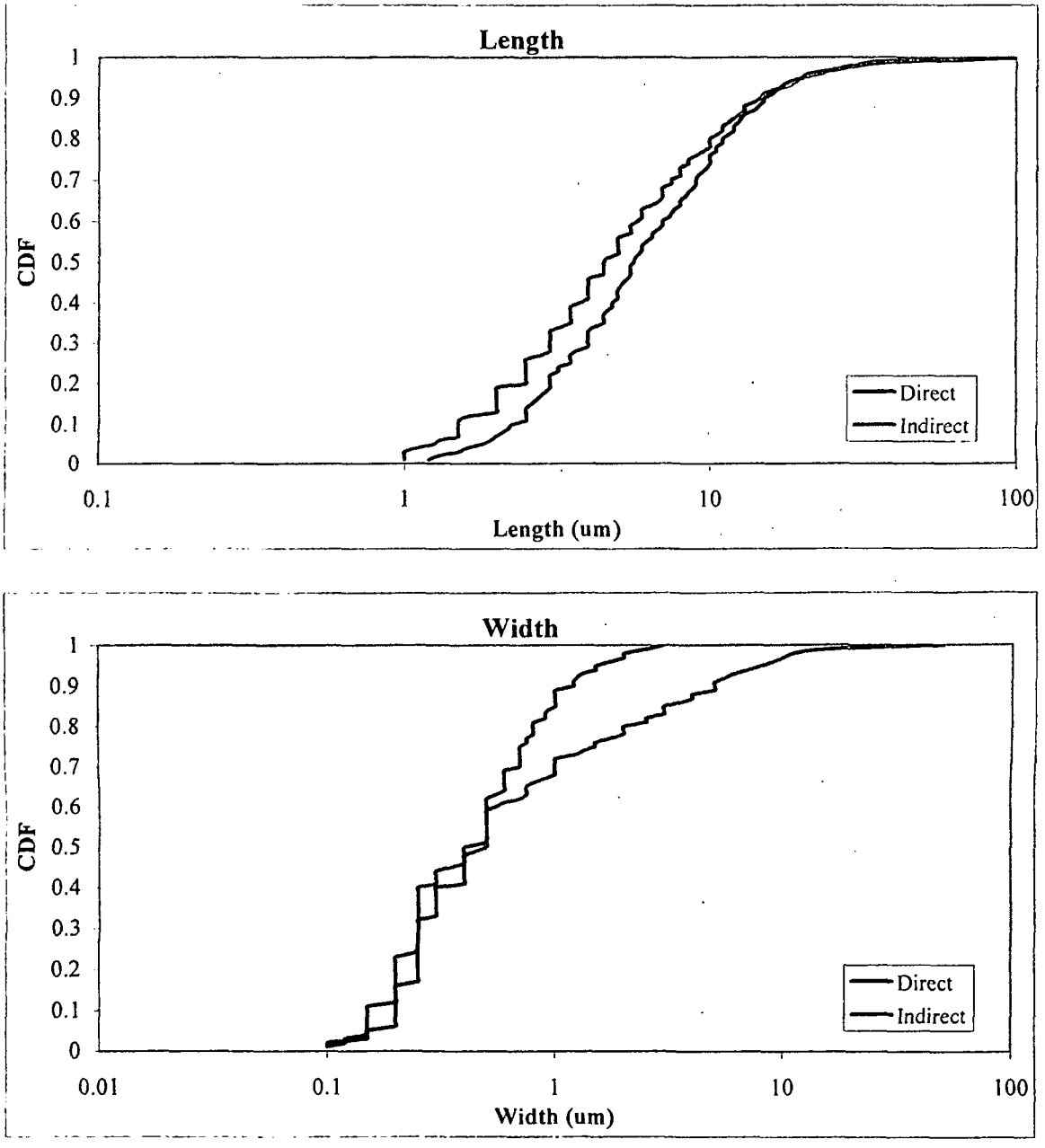


FIGURE 2
LA PARTICLE SIZE DISTRIBUTIONS FOR
31 PAIRED DIRECT AND INDIRECT PREPARATIONS



Graphs are based on paired results for the 31 samples analyzed by both direct and indirect preparation methods [N=754 structures in direct analysis, 1,617 structures by indirect analysis].

FIGURE 3
EFFECT OF ALTERNATIVE SURROGATE VALUES FOR NON-DETECTS ON THE EXPECTED SAMPLE MEAN FOR ASBESTOS

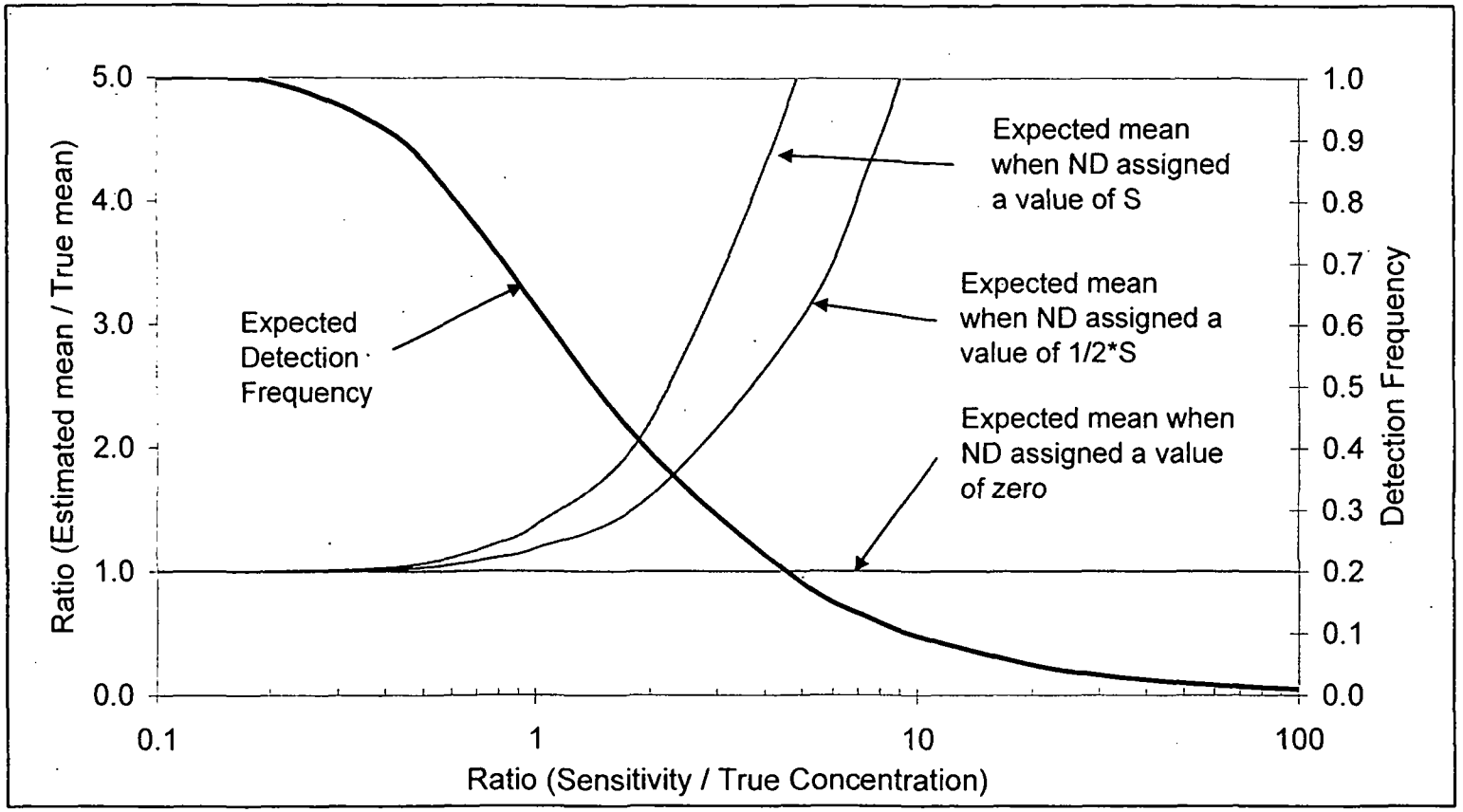


FIGURE 4. COUNTING UNCERTAINTY IN SINGLE SAMPLES

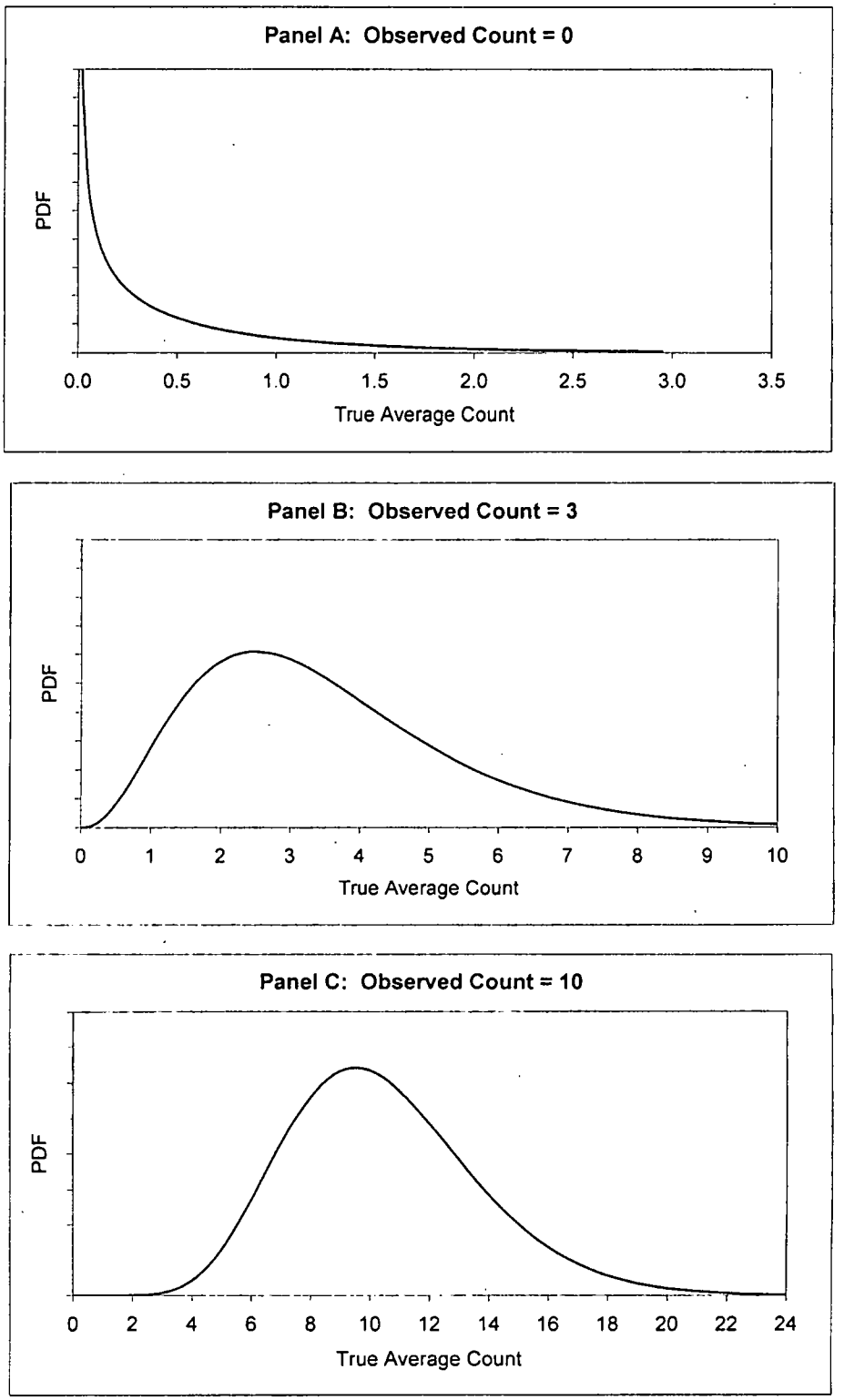
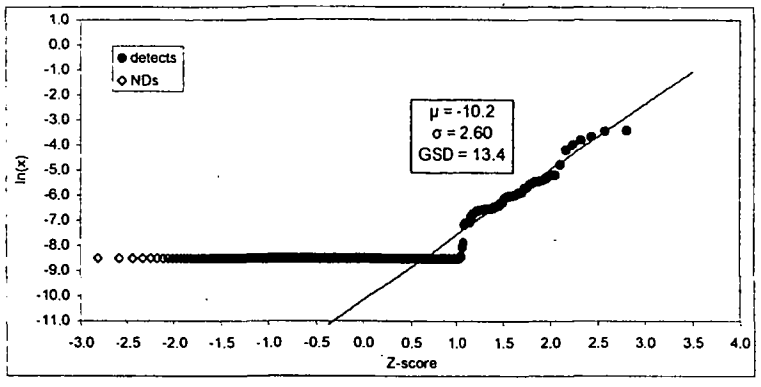
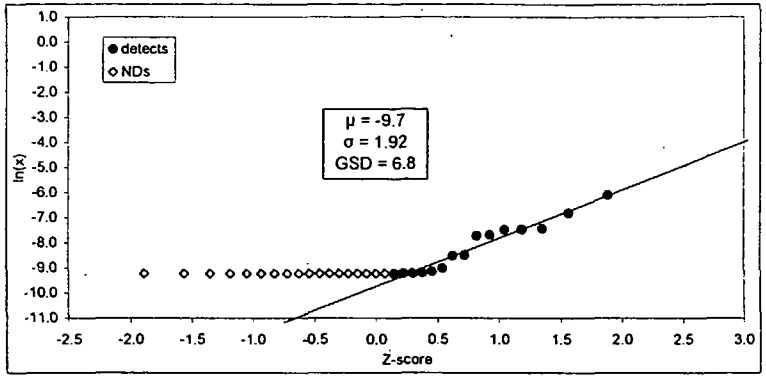


FIGURE 5. LOG-PROBABILITY PLOTS OF EXAMPLE AIR DATA SETS

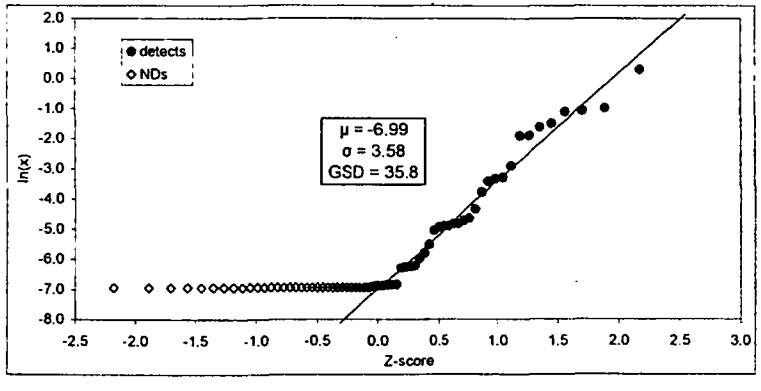
Panel A: Ambient Air (All)



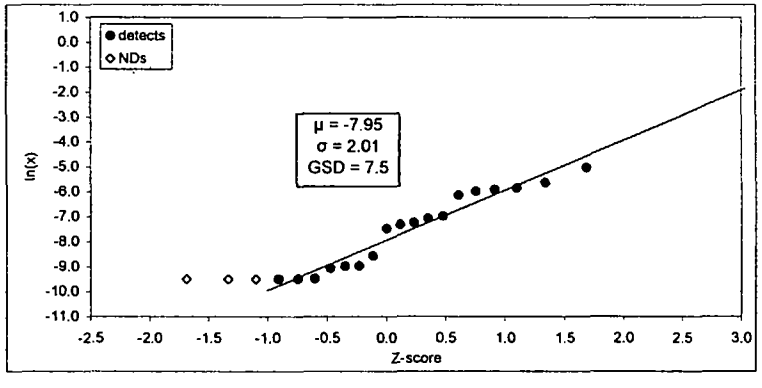
Panel B: Ambient Air (33 reanalyzed)

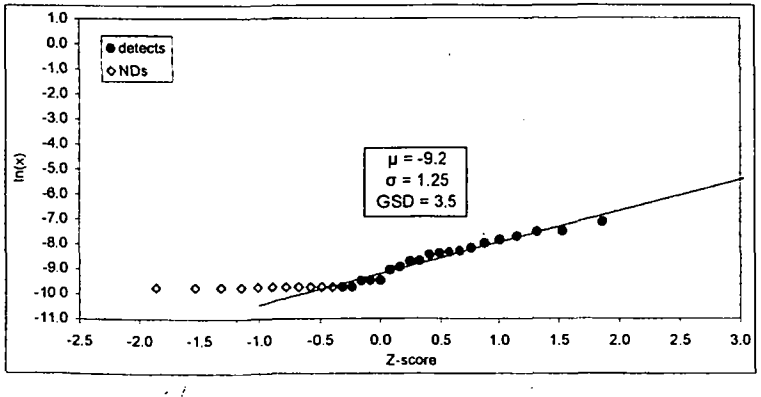


Panel C: Activity-Based Samples



Panel D: Personal Indoor Air





NOTES:
Detects are shown as solid red diamonds
Non-detects are shown as open blue diamonds
Non-detects are plotted at an aribtrary value equal to the lowest detect
Lognormal parameters (μ, σ) are based on detects only

FIGURE 6
EFFECT OF SAMPLING VARIABILITY AND MEASSUREMENT ERROR ON CHEBYCHEV INEQUALITY UCL VALUES

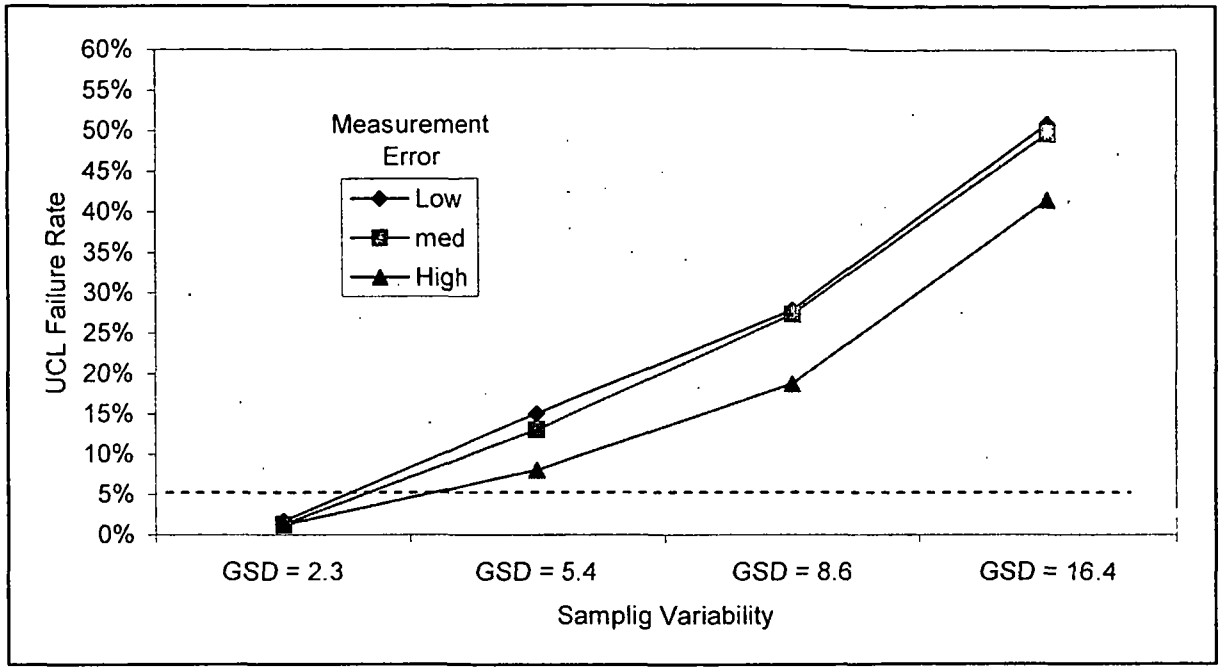


FIGURE 7
UCL COVERAGE FOR THE PLN-LAND$_{95}$ METHOD

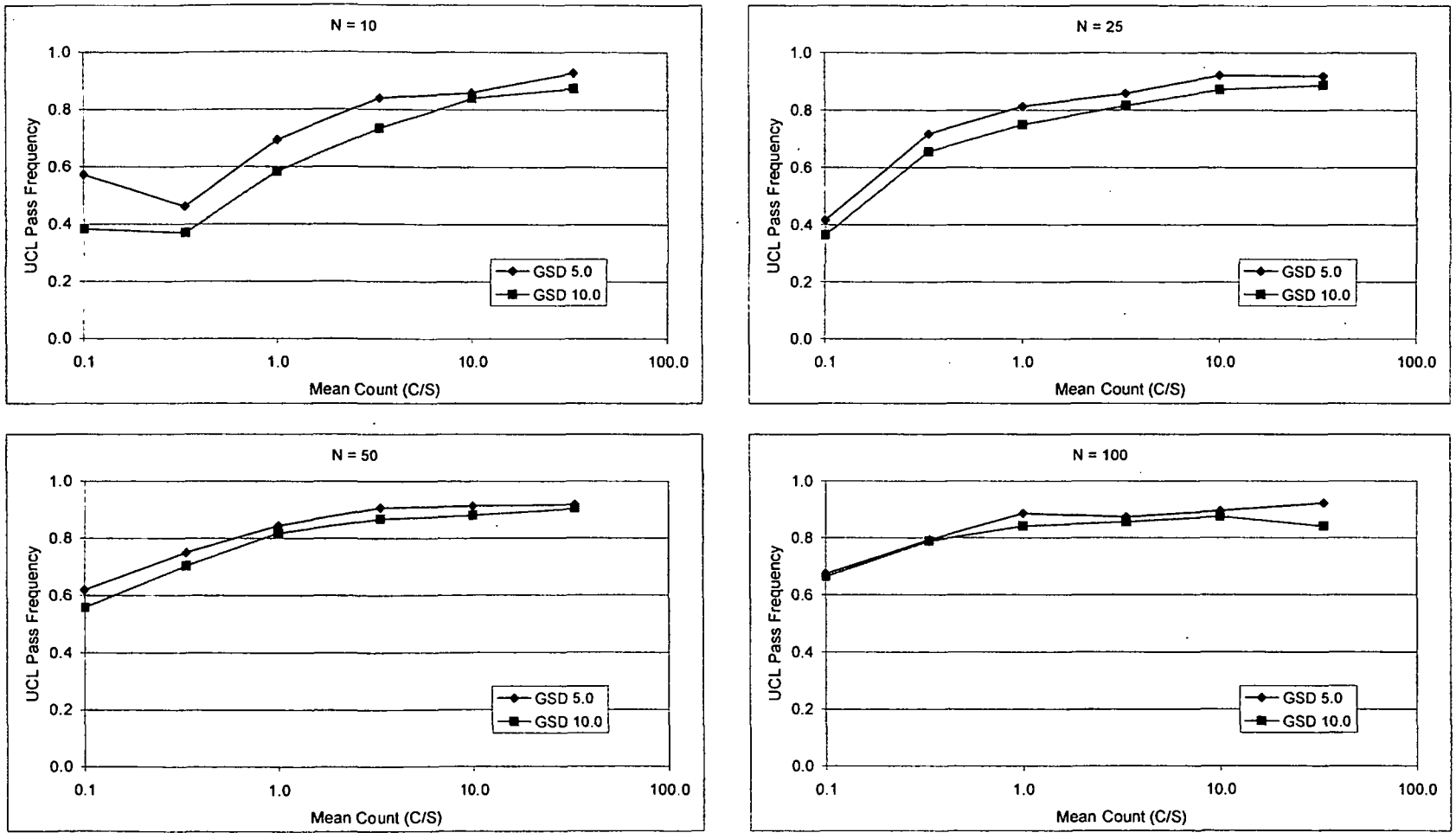


FIGURE 8
UCL COVERAGE FOR THE PLN-LAND$_{97.5}$ METHOD

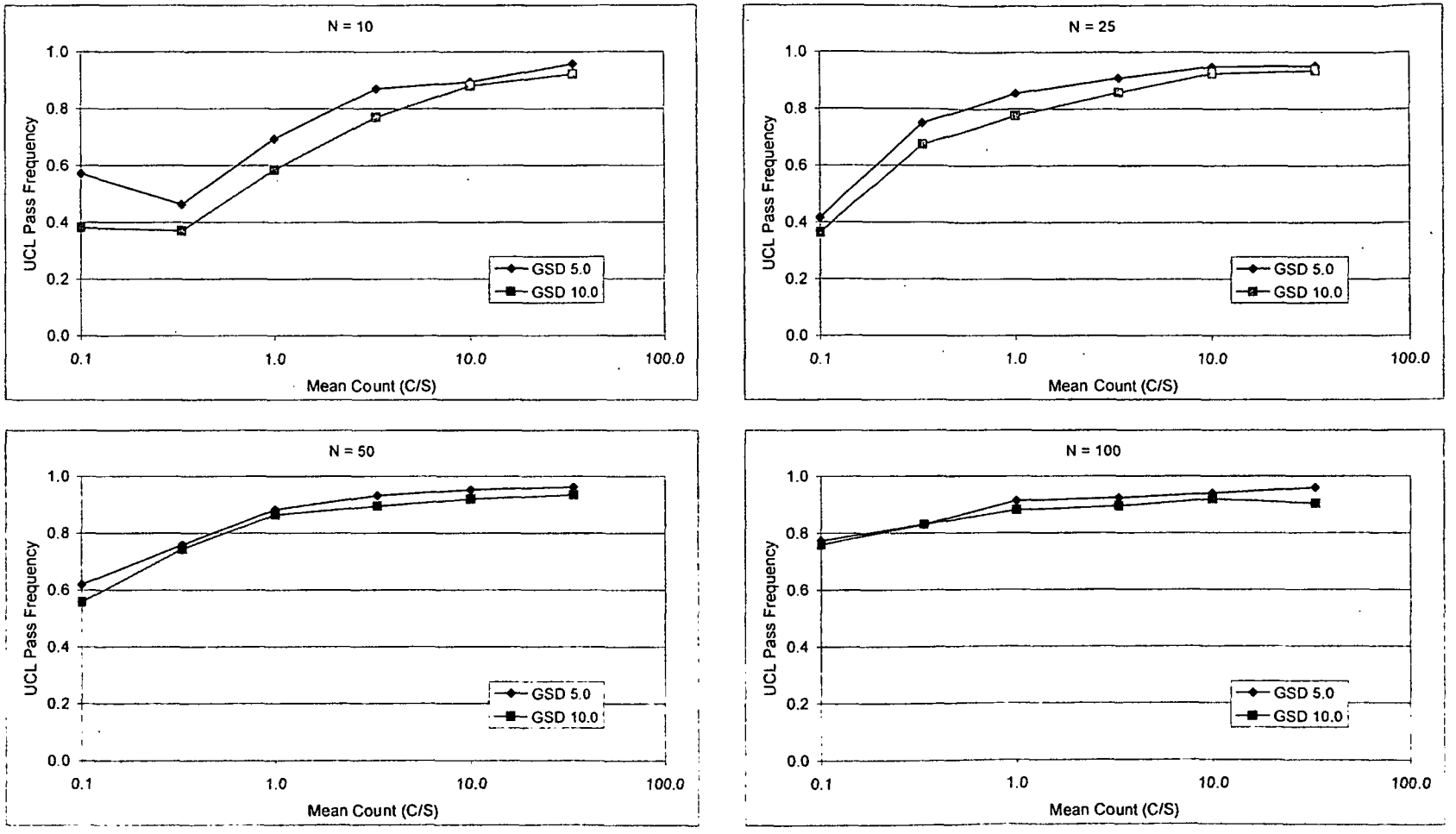


FIGURE 9
UCL COVERAGE FOR THE TWO-STEP BOOTSTRAP METHOD

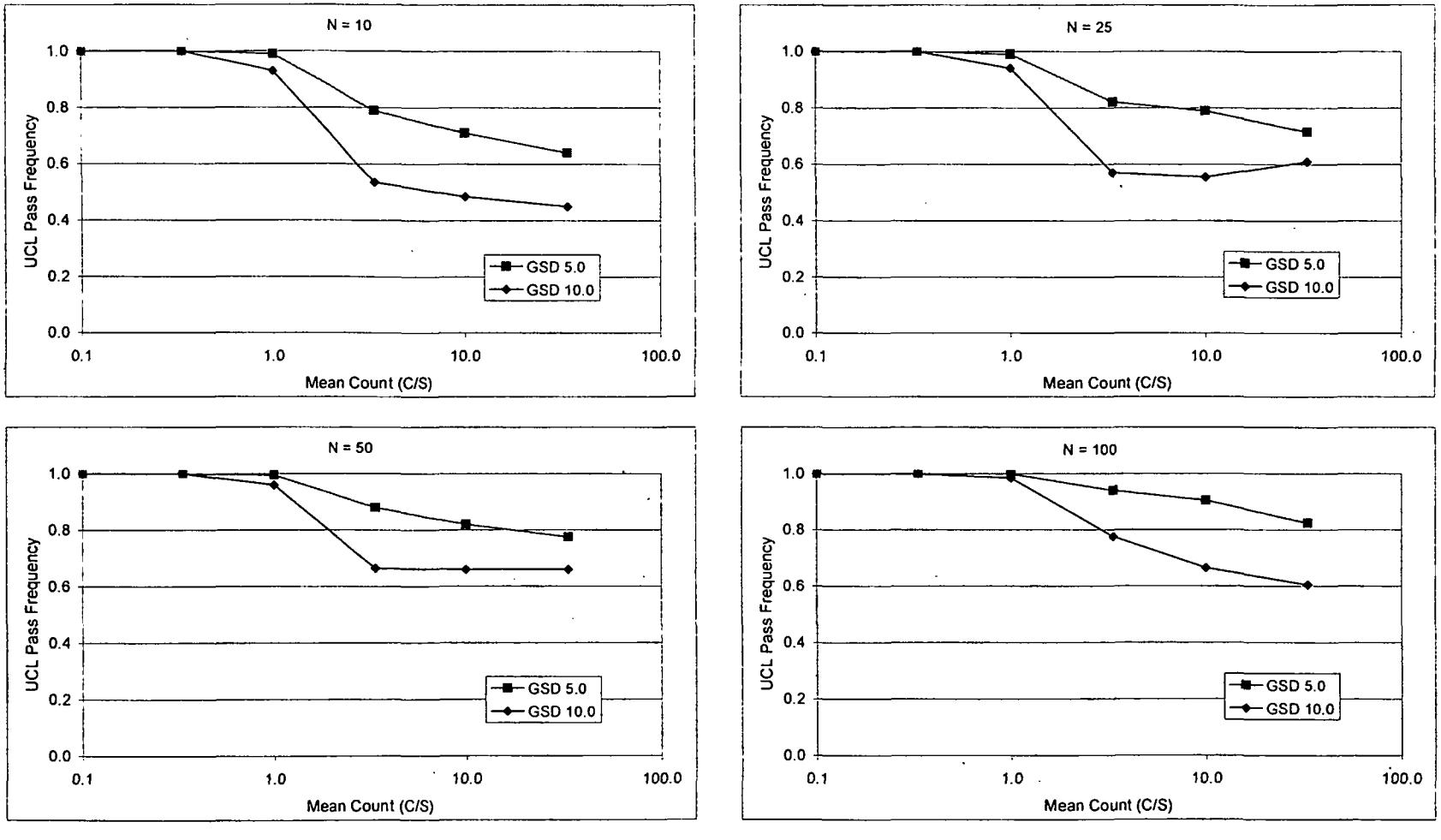


FIGURE 10

ATTACHMENT 1

STATISTICAL COMPARISON OF TWO POISSON RATES BASED ON THE RATIO IN RATES

1.0 Basic Equations

The equations presented below are from Nelson (1982). The purpose of the statistical test is to compare two Poisson rates:

Rate 1 = Y / t
Rate 2 = X / s

where

Y and X = The number of "hits" (e.g., the numbers of particles counted)
t and s = The "size" of the observations (e.g., the effective volume of air examined)

The test is based on evaluating the uncertainty bounds around the ratio of the rates:

ρ = Rate 1 / Rate 2 = (Y/t) / (X/s)

Let γ represent the confidence interval around the ratio. Limits on the two-sided $100\gamma\%$ confidence interval for ρ are:

ρ(lower bound) = {(Y/t)/[(X+1)/s]} / F[(1+γ)/2; 2X+2, 2Y]
ρ(upper bound) = {[(Y+1)/t]/(X/s)} · F[(1+γ)/2;2 Y+2, 2X]

where:

F[x;df1,df2] = F distribution with df1 and df2 degrees of freedom

2.0 Interpretation

- If the confidence interval includes the value 1, the data are consistent with the hypothesis that Rate 1 and Rate 2 are not different from each other at the [100(1-γ)]% significance level.

- If the lower bound on the confidence interval around ρ exceeds 1, then Rate 1 is greater than Rate 2 at the $[100(1-\gamma)]\%$ significance level
- If the upper bound on the confidence interval around ρ is less than 1, then Rate 2 is greater than Rate 1 at the $[100(1-\gamma)]\%$ significance level

3.0 References

Nelson W. 1982. Applied Life Data Analysis. John Wiley & Sons, New York.

ATTACHMENT 2

MONTE CARLO SIMULATION OF MEASUREMENT ERROR UNCERTAINTY IN THE MEAN OF MULTIPLE SAMOPLE

See Microsoft Excel Spreadsheet entitled
"Uncertainty in the Mean of 5 Data Sets.xls"

ATTACHMENT 3

FITTING A DATA SET TO A POISSON-LOGNORMAL MODEL

ATTACHMENT 3

FITTING A DATA SET TO A POISSON-LOGNORMAL MODEL

Basic Model

Let C_k represent the concentration of LA in a sample of air (s/cc) collected from some specified location. Assume that a set of n individual samples collected from this location have concentration values that are distributed lognormally:

$$C_k \sim LN(\mu, \sigma)$$

The true value of the concentration of LA in any random sample C_k can not be measured exactly, but can only estimated by microscopic analysis. This is because the number of structures x_k observed in an analysis of some volume V_k from that sample is a random Poisson variable, given by:

$$x_k \sim Poisson[C_k \cdot V_k)]$$

Poisson Lognormal (PLN) Distribution Function

The situation described above may be characterized by a mixing distribution that captures both the large scale lognormal variability and the within-sample Poisson variability (Haas et al. 1999). The probability density for the Poisson-lognormal model (PLN) is:

$$Prob(x \mid V, \mu, \sigma) = \int_0^{\infty} \frac{(CV)^x \exp(-CV)}{x!} \frac{1}{C\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\ln C - \mu}{\sigma}\right)^2\right] dC$$

where:

C = concentration (LA s/cc)
V = volume of sample analyzed (cc)
x = number of LA structures observed (0, 1, 2, 3, ...)
μ = log-mean of the distribution of C
σ = log-standard deviation of the distribution of C

Figure A3-1 illustrates the shape of the probability density curve for three example cases where the arithmetic mean concentration is 0.001 s/cc, the volume of air evaluated is 1,000 cc (this equals a sensitivity of 0.001 cc^{-1}), and the variability corresponds to GSD values of 3, 5, or 10.

Evaluation of the Poisson-Lognormal Probability Density Function.

To evaluate the Poisson-lognormal density function, Haas et. al. (1999) recommended re-expressing the integrand in a form suitable for Gauss-Hermite numerical quadrature. However, in testing this approach for use at the Libby Site, it was discovered that for many combinations of $\{x;V,\mu,\sigma\}$, Gauss-Hermite quadrature could not dependably provide the precision necessary for this analysis. Therefore, for the purposes of use at the Libby Site, the PLN probability density was evaluated by dividing the integral into two parts (0 to 100 and 100 to infinity), and using a combination of *IMSL* numerical quadrature routines *(int_fcn* for the first part and *int_fcn_inf* for the second part). Both routines are based on a globally adaptive technique using Gauss-Kronrod rules. To minimize scaling issues in the numerical integration, concentration values were expressed in units of s/L rather than s/cc.

Maximization of the Poisson-Lognormal Log-Likelihood

Fitting the Poisson-lognormal distribution to a set of n paired observations of count (x_k) and volume (V_k), designated by the vectors (**x,V**), requires maximization of the sample log-likelihood:

$$\mathcal{L}(\mathbf{x}|\mathbf{V},\mu,\sigma) = \sum_{k=1}^{n} \ln Prob(x_k; v_k, \mu, \sigma)$$

which was accomplished using the *IMSL* quasi-Newton routine *min_uncon_multivar*.

Figure A3-2 provides a graphical display of the PLN fit to an example data set (purely hypothetical) of 100 samples (provided in Table A3-1). In this example, the fitted parameters were $\mu = 1.328$ and $\sigma = 1.2399$.

References:

Haas CN, Rose JB, Gerba CP. 1999. Quantitative Microbial Risk Assessment. John Wiley and Sons, New York.

IMSL C Functions for Scientific Programming, Version 1.0. Visual Numerics, Houston, Texas. http://www.vni.com

FIGURE A3-1

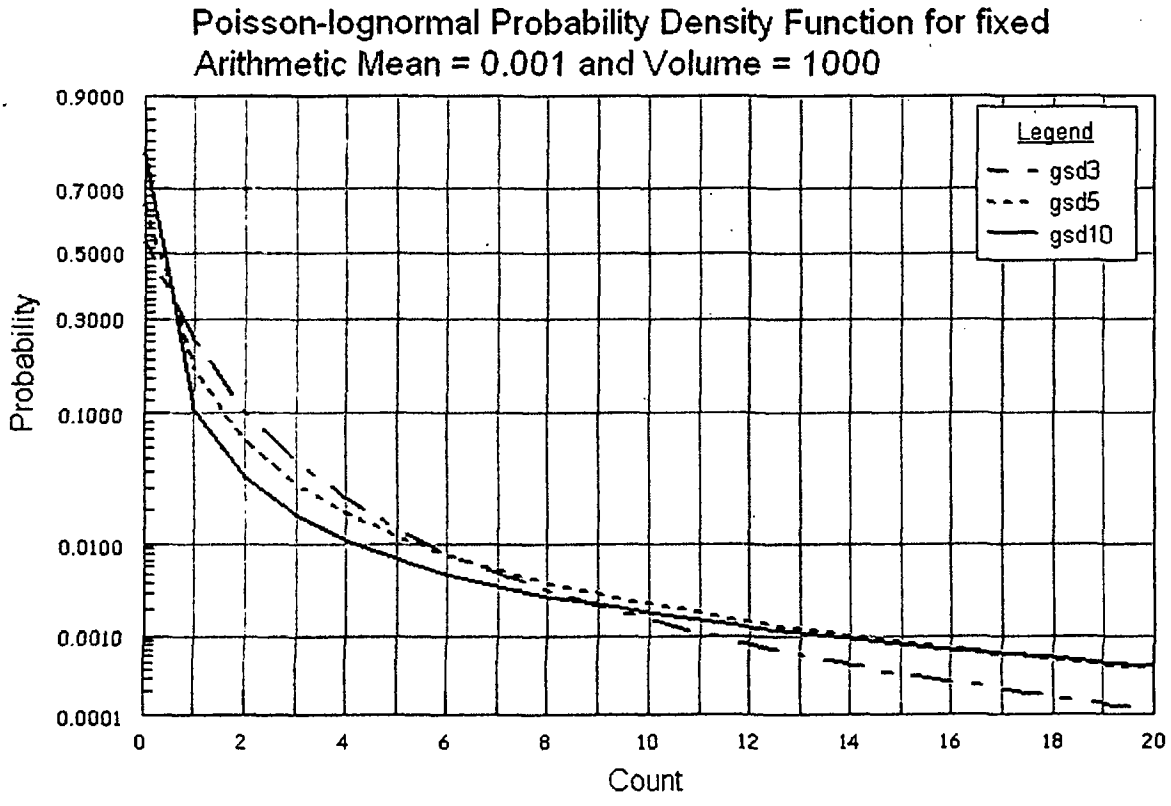

Poisson-lognormal Probability Density Function for fixed
Arithmetic Mean = 0.001 and Volume = 1000
Probability
0.9000
0.7000
0.5000
0.3000
0.1000
0.0100
0.0010
0.0001
0
2
4
6
8
10
12
14
16
18
20
Count
Legend
gsd3
gsd5
gsd10

FIGURE A3-2
FIT OF THE PLN DISTRIBUTION TO AN EXAMPLE DATA SET

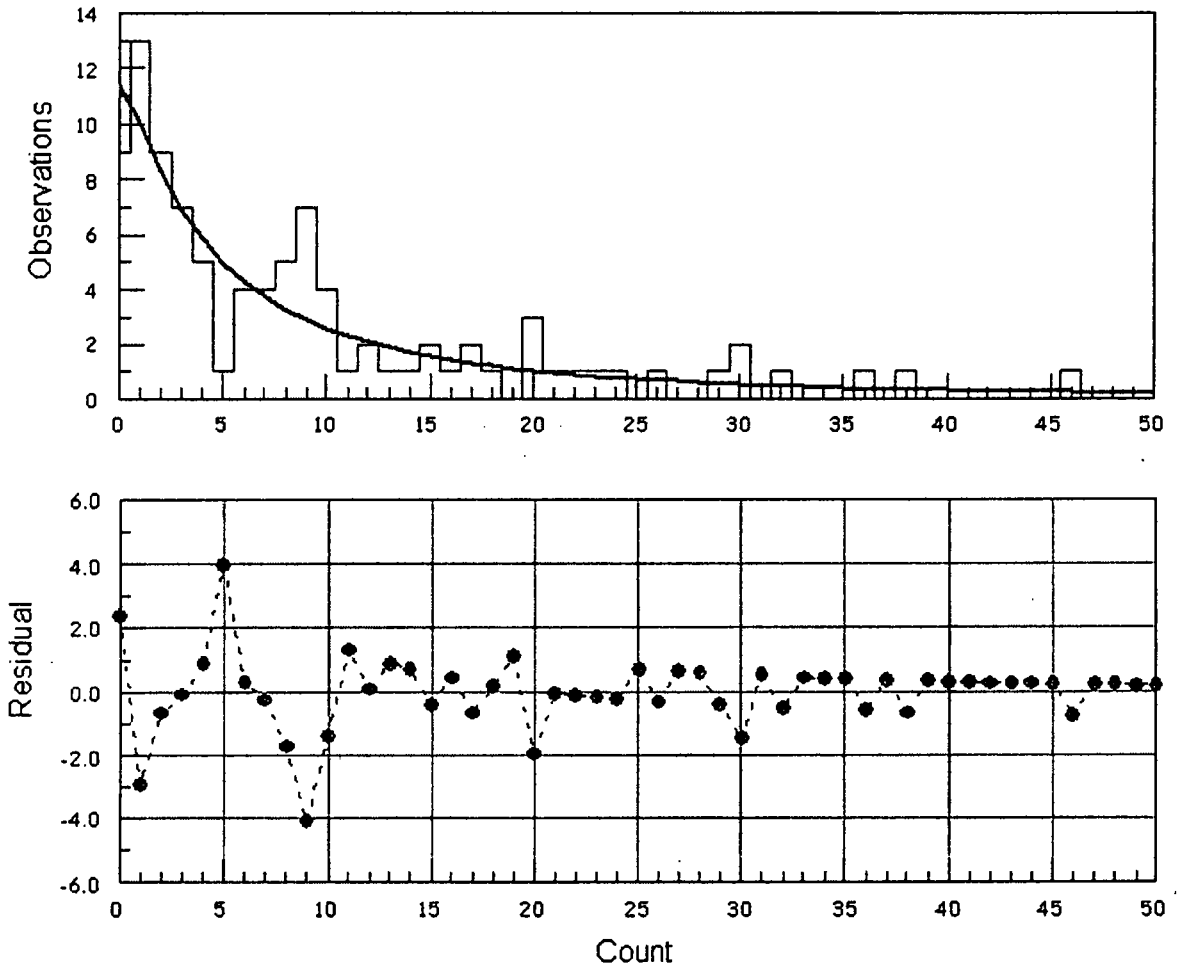


TABLE A3-1
EXAMPLE DATA USED IN FIGURE A3-2

Sample	Count	Volume (L)	Sample	Count	Volume (L)
1	9	3.14	51	2	1.62
2	2	2.70	52	26	4.00
3	2	2.08	53	10	3.94
4	3	2.00	54	1	2.76
5	4	0.72	55	3	1.92
6	1	1.16	56	1	1.52
7	16	2.70	57	20	2.56
8	5	0.46	58	1	0.62
9	12	2.68	59	78	1.72
10	1	0.98	60	11	2.38
11	9	0.76	61	8	0.80
12	30	1.90	62	6	2.82
13	7	2.08	63	1	1.90
14	9	2.58	64	56	3.94
15	3	1.66	65	29	2.10
16	2	3.32	66	6	3.80
17	4	0.58	67	2	0.64
18	0	2.50	68	3	3.82
19	8	2.70	69	4	2.32
20	10	1.68	70	9	1.50
21	1	0.18	71	0	1.48
22	21	2.62	72	6	3.46
23	15	3.82	73	46	2.58
24	1	0.22	74	0	0.12
25	32	2.42	75	53	2.52
26	22	1.90	76	10	1.28
27	6	2.70	77	8	0.58
28	537	3.10	78	7	2.32
29	1	1.14	79	30	2.92
30	20	1.34	80	0	0.24
31	15	1.66	81	2	0.74
32	2	3.72	82	1	0.24
33	3	2.00	83	3	1.54
34	3	2.68	84	1	0.12
35	0	3.00	85	7	1.76
36	0	1.40	86	9	1.54
37	17	3.16	87	71	3.38
38	9	2.72	88	14	3.66
39	4	2.36	89	8	3.46
40	7	1.26	90	10	2.14
41	0	0.26	91	1	0.86
42	20	0.76	92	8	0.10
43	17	2.62	93	13	2.38
44	0	1.82	94	1	0.84
45	4	0.78	95	127	2.84
46	23	3.76	96	2	0.36
47	36	1.46	97	2	0.34
48	24	1.14	98	18	1.38
49	9	2.40	99	12	3.80
50	38	3.56	100	0	0.56

APPENDIX A

1,837 Samples Analyzed by ISO 10312 and AHERA

See Microsoft Excel Spreadsheet "AppA_AHERA vs ISO.xls"

APPENDIX B

16 Air Samples Evaluated by Direct and Indirect Analysis

See Microsoft Excel Spreadsheet "AppB_I vs D_Original 16.xls"

APPENDIX C

Study Design and Selection of 31 Air Samples for Re-Analysis Using Indirect Preparation

APPENDIX D

Detailed Results for 31 Air Samples Analyzed By Both Direct and Indirect Preparation

See Microsoft Excel Spreadsheet "AppD_I vs D Pilot Results.xls"